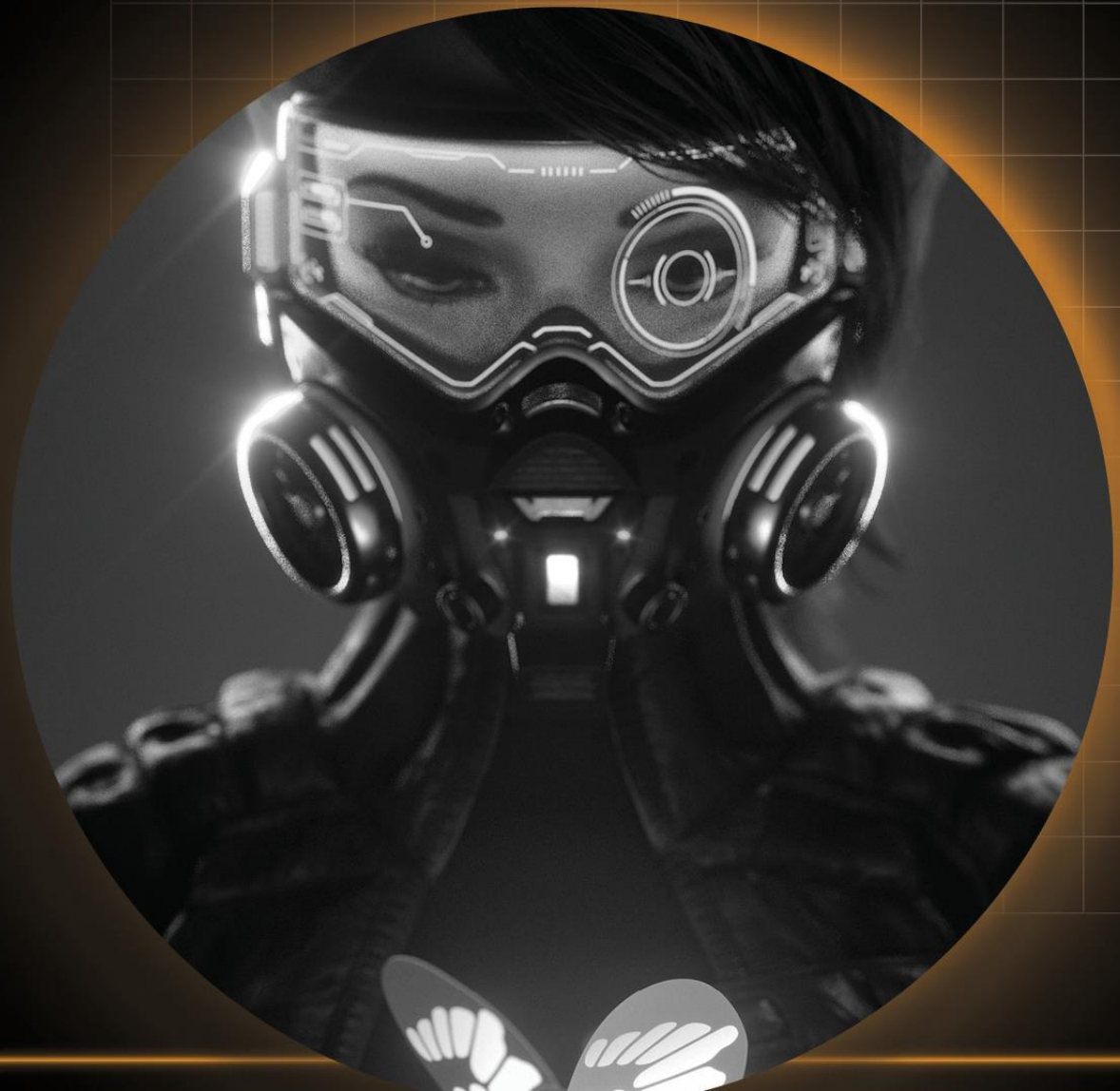
EVOLUTION STRATEGIES ALGORITHMS AS A SCALABLE ALTERNATIVE TO REINFORCEMENT LEARNING

Emmanuelle Rodrigues Nunes
September 2021



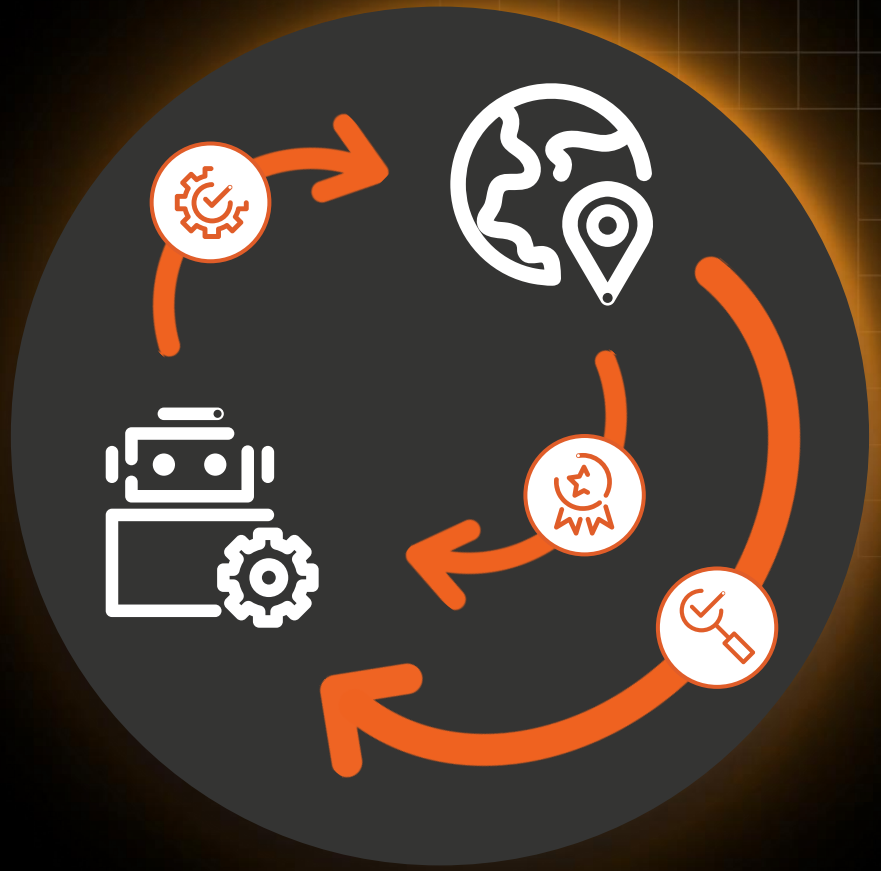
ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

LIONBRIDGE

REINFORCEMENT LEARNING

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

- Uses gradient descent
- Needs continuous function



LIONBRIDGE

EVOLUTION ALGORITHMS

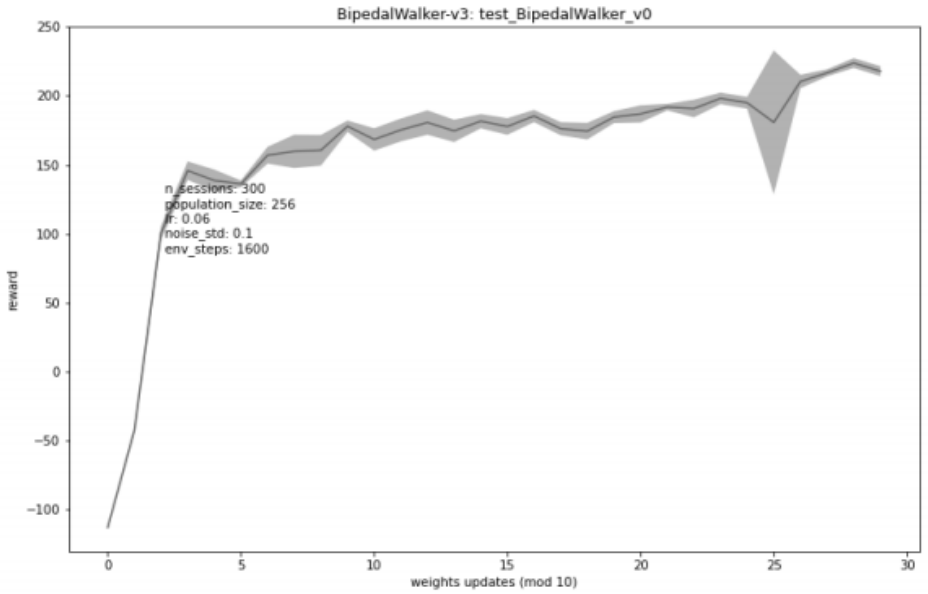
ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

- Black-box optimisation algorithm derived from the Evolutionary Algorithms' family (EA).
- The algorithm starts with an individual who suffers a mutation. The fittest individual from the parent and the offspring becomes the next generation's parent.
- Evolution Strategies (ES) was initially inspired by natural selection.
- ES is capable of self-adaptation.

LIONBRIDGE

EXPERIMENT

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS



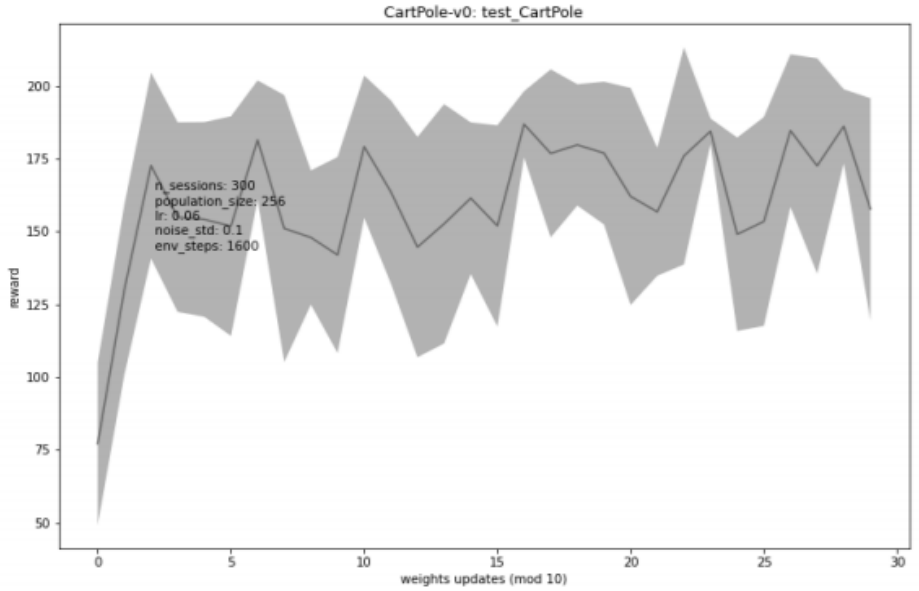


Table 1. Experiment configuration

Configuration	Value
Sessions	300
Environment Steps	1600
Population Size	256
Learning Rate	0.06
Noise Standard Deviation	0.1
Noise Decay	0.99
Learning Rate Decay	1
Decay Steps	20
Evaluation Step	10
Hidden Sizes	(40, 40)

LIONBRIDGE

DISCUSSION

THE GREAT BENEFIT OF THIS APPROACH IS THAT IT IS PARALLELISABLE WITH LITTLE EFFORT.

- ES only requires the fitness scores and high-level parameter distribution information.
  - The cost is treating the network as a black box.
  - In situations with only sparse rewards, turning the problem from a ”mostly black box” into an ”entirely black box” is worth the performance improvements.
- The best solution is to combine both methods.
  - For reactive policies or sparse rewards situations, ES is a good candidate.
  - Gradient-based methods using RL are good models when a rich feedback signal is available.

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Thank you.

EMMANUELLE RODRIGUES NUNES

LIONBRIDGE